Supplementary Materials and Methods

Table of Contents

1 Patients and control subjects for SNP genotyping

1.1 Neuroblastoma patients in discovery case series

The neuroblastoma patients in the study were children diagnosed with neuroblastoma or ganglioneuroblastoma and registered through the Children's Oncology Group (COG). The blood samples from the neuroblastoma cases were identified through the COG Neuroblastoma bio-repository for specimen collection at the time of diagnosis. The majority of specimens were annotated with clinical and genomic information (see the table below). A subset of the samples have been assigned into one of three risk groups (low-risk, intermediate-risk and high-risk) based on analysis of well-defined prognostic factors, including patient age at diagnosis[1], International Neuroblastoma Staging System (INSS) stage [2], tumor histopathology[3,4], DNA index[5], and *MYCN* amplification status[6,7].

Clinical characteristics of 1529 patients with information from COG.

Characteristic	Number (%)
Age	
< 1 yr	548 (36%)
≥ 1 yr	981 (64%)
INSS Stage	
1,2,3,4s	828 (55%)
4	675 (45%)
Unknown	26
MYCN	
Not Amplified	1153 (82%)
Amplified	261 (18%)
Unknown	115
DNA Index	
Hyperdiploid	884 (64%)
Diploid	487 (36%)
Unknown	158
Histology	
Favorable	680 (56%)
Unfavorable	543 (44%)
Unknown	306
Risk	
Low/Intermediate	815 (55%)
High	659 (45%)
Unknown	55

Eligibility criterion for genotyping was availability of 1.5 μg of high quality DNA from a tumor-free source such as peripheral blood or uninvolved (with tumor) bone marrow mononuclear cells. Because neuroblastoma in the United States is demographically a disease of Caucasians of European descent [8], we limited our analyses to this ethnicity group to minimize genetic heterogeneity, as outlined in following sections.

Previously published GWAS on neuroblastoma interrogated a subset of the sample collection used in the current study. Specifically, the Maris et al study [9] analyzed 1032 blood samples genotyped on the HumanHap550 arrays, yet the Capasso et al study [10] analyzed a subset of 397 high-risk cases from the Maris et al study. All the genotypes for patients used in the current study will be deposited into dbGAP (http://www.ncbi.nlm.nih.gov/gap).

1.2 Control subjects

The control group included 3,254 children of Caucasian ancestry who were recruited and genotyped by the Center for Applied Genomics at The Children's Hospital of Philadelphia (CHOP) and who passed the stringent quality control procedures discussed below. The controls were recruited from multiple sites within the CHOP Health Care Network, including four primary care clinics and several group practices and outpatient practices that included well child visits. Consenting was performed by nursing and medical assistant staff under the direction of CHOP clinicians.

Eligibility criteria for control subjects were: 1) self-reported as Caucasian; 2) availability of 1.5 μg of high quality DNA from peripheral blood mononuclear cells; and 3) no known medical disorder, including cancer, based on self-reported intake questionnaire or clinician-based assessment. The Research Ethics Board of CHOP approved the study,

and written informed consent was obtained from all subjects. All DNA samples were extracted from whole blood. Although these control subjects were all self-identified Caucasians, we used IBS clustering algorithm implemented in PLINK[11] to infer a homogeneous group of control subjects of European ancestry for association analysis.

All control subjects in the discovery cohort were genotyped using the Illumina HumanHap550 array platform. The control subjects in the replication cohort were recently recruited at the Children's Hospital of Philadelphia and were genotyped on the new generation of Illumina Human610-Quad array that include both SNP and CNV markers.

2 Whole-genome SNP genotyping

SNP genotyping was performed using the Illumina Infinium™ II BeadChip (Illumina, San Diego, CA, USA) [12,13] according to methods detailed elsewhere[14], and summarized below.

DNA samples were surveyed for quality both by optical density spectrophotometry and the pico-green assay at the Center for Applied Genomics at the Children's Hospital of Philadelphia. A total of 750 nanograms of genomic DNA was used to genotype each sample, according to the manufacturer's guidelines. On day one, genomic DNA was amplified 1000-1500 fold; and on day two, the amplified DNA was fragmented to ~300-600 basepairs (bp), precipitated and resuspended followed by hybridization onto a BeadChip. Single base extension (SBE) utilized a single probe sequence of approximately 50 bp long designed to hybridize immediately adjacent to the single nucleotide polymorphism (SNP) query site. Following targeted hybridization to the bead

array, the arrayed SNP locus-specific primers (attached to beads) were extended with a single hapten-labeled dideoxynucleotide in the single base extension reaction. The haptens were subsequently detected by a multi-layer immunohistochemical sandwich assay. The Illumina BeadArray Reader scanned each BeadChip at two wavelengths and created an image file. As BeadChip images were collected, intensity values were determined for all instances of each bead type, and data files were created that summarized intensity values for each bead type. These files consisted of intensity data that was loaded directly into Illumina's genotype analysis software, BeadStudio. A bead pool manifest created from the Laboratory Information Management System (LIMS) database containing all the BeadChip data was loaded into BeadStudio along with the intensity data for the samples. BeadStudio used a normalization algorithm to minimize BeadChip to BeadChip variability. Once the normalization was complete, the clustering algorithm was run to evaluate cluster positions for each locus, and then assign individual genotypes. All SNP genotype calls were based on the default genotype clustering file provided by Illumina for the corresponding genotype platforms.

3 SNP association analysis

Below we describe in detail the quality control procedure utilized for the SNP association analysis in GWAS for the discovery case series. The quality control procedure for the replication cohort is largely similar to those performed on the discovery cohort.

3.1 Overlap of the HumanHap550 v1 and v3 arrays

Since a portion of the individuals in the discovery cohort are genotyped by the HumanHap550 v1 array (n=859) while others are genotyped by the v3 array, our analysis only concerns the markers shared by the v1 and v3 array. The HumanHap550

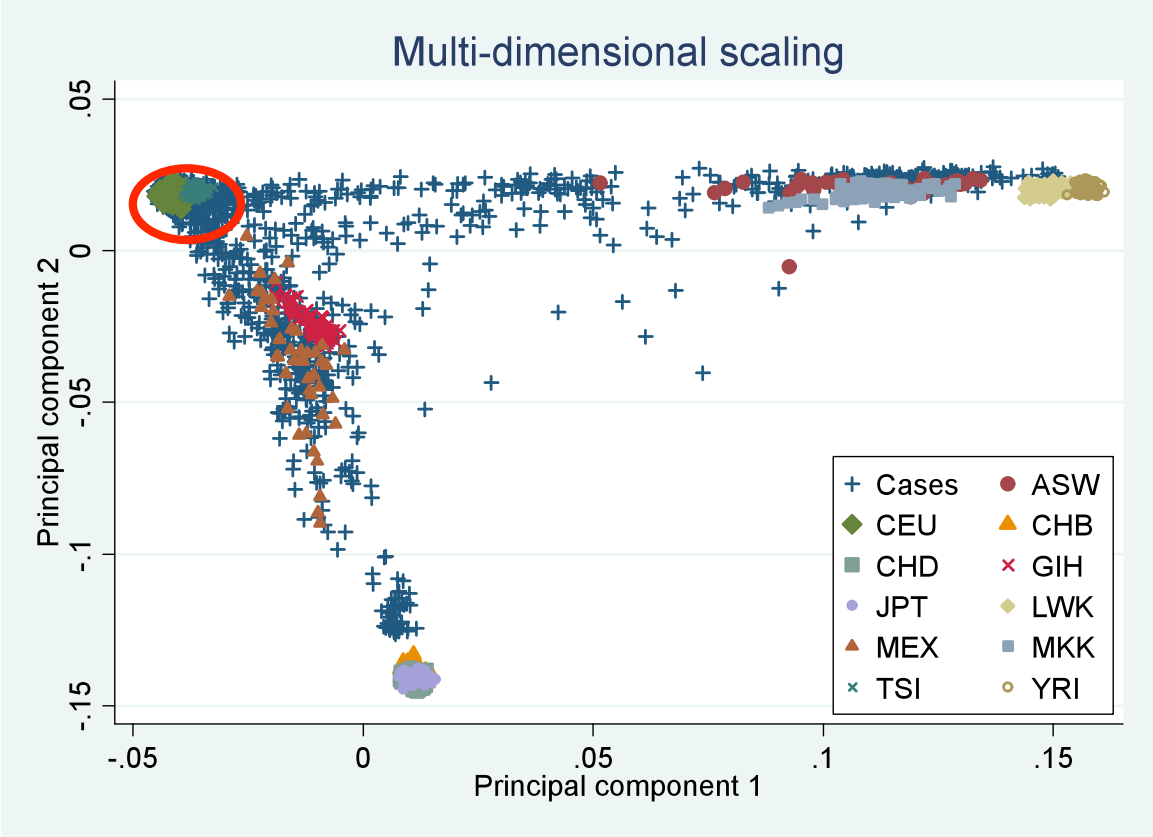
v1 array contains 555,175 markers, while the v3 array contains 561,288 markers, including 544,902 markers that are shared by the two arrays.

3.2 Low genotype call rate (<95%)

The call rate is calculated based on the number of “No Call” genotypes with default genotyping calling algorithm as implemented in the Illumina BeadStudio software. The call rate per individual was assessed by the PLINK software. A total 96 cases were excluded from analysis due to low call rate (<95%)

3.3 Inferring individuals of European ancestry

We used Multi-Dimensional Scaling (MDS), as implemented in the PLINK software, for inferring population structure in the neuroblastoma data set. Comparing self-identified ancestry with the MDS-inferred ancestry confirmed the reliability of MDS to identify genetically inferred individuals of European ancestry. In total, 1642 neuroblastoma patients of European ancestry are clustered towards the upper left side of the triangle (red circle) compared with 11 HapMap3 populations, and defined by Principal component 1 (X-axis) less than -0.02, and Principal component 2 (Y-axis) greater than 0.



3.4 Detection and elimination of cryptic relatedness and duplicated genotyping

We have calculated genome-wide IBS estimates for all pairwise comparisons among all case subjects and control subjects. To detect cryptic relatedness and potential duplicated genotyping within our data sets, we have applied a two-step procedure to calculate pairwise IBD estimates between all individuals. First, we examined MDS and only keep in our data sets those individuals of inferred European ancestry (see description in previous section), with call rates greater than 95%; second, we re-calculated genome-wide IBS estimates and re-calculate the IBD estimates among the remaining individuals of European ancestry using the PLINK software. This two-step procedure ensures that allele frequency differences between populations do not lead to biases in IBD estimations. We applied a stringent threshold for detecting cryptic

relatedness: any pairs of subjects with IBD>0.15 were processed such that only unrelated subject were remained in the final association test. In total, 15 neuroblastoma patients were excluded.

3.5 Low call rate per marker (< 95%)

Markers with call rate less than 95% were excluded from analysis. The call rates were calculated by the PLINK software. A total of 8,281 markers were excluded from association analysis in this step.

3.6 Minor Allele Frequency (individuals of European ancestry)

Markers with Minor Allele Frequency (MAF) less than 5% were excluded from our analysis. The MAF are calculated by the PLINK software. A total of 50,869 markers were excluded from association analysis in this step.

3.7 Hardy-Weinberg Equilibrium (individuals of European ancestry)

Markers with Hardy-Weinberg Equilibrium P-value less than 0.001 were excluded from analysis. A total of 5,415 markers were excluded from association analysis in this step.

3.8 Matching controls

Based on genome-wide IBS estimates for all pairwise comparisons among all case subjects and control subjects, we identified two matched controls for each case individual to correct the potential effects of population structure.

3.9 Population stratification

To further address the concerns on population stratification, we have also applied EigenStrat software [15] to re-perform all association tests on the case and control subjects passing the QC threshold above. The P-values for the SNPs reported are very similar with those obtained from standard allelic chi-square test, further implicating the

effectiveness of MDS approach in removing population outliers. The most significant marker rs110419 has an EigenStrat P-value of 2.35×10^{-10}, which is indeed slightly lower than the unadjusted P-value (5.12×10^{-10}).

3.10 Final counts of subjects and markers passing QC

Applying the QC measures mentioned in all the previous sections, we were left with 1,627 cases for the discovery series, with 3,254 matching controls and 480,279 SNPs for association analysis.

3.11 Genotype imputation on whole-genome SNP data

Genotype imputation was performed using the MACH software (http://www.sph.umich.edu/csg/abecasis/MaCH/index.html) on both the discovery and replication case series. The software version 1.0.16 was used in the study, and the default two-step procedure was adopted for imputation. The HapMap phased haplotypes (release 22) on CEU subjects, as downloaded from the HapMap database (http://www.hapmap.org), were used as input files for SNPs and phased haplotypes.

4 Somatic copy number alterations in tumor samples

4.1 Tumor samples

The tumor samples include a total of 701 tumors which were genotyped by the Illumina whole-genome SNP genotyping array, most of whom have matched blood samples that were also genotyped by the same arrays. However, seven cases were eliminated from the clinical correlative analyses because of missing data. The clinical characteristics for the samples are summarized below:

Clinical characteristics of 694 patients with available information

Characteristic	Number (%)

Age	
< 1 yr	168 (24%)
≥ 1 yr	526 (76%)
INSS Stage	
1,2,3,4s	247 (36%)
4	445 (64%)
Unknown	2
MYCN	
Not Amplified	519 (76%)
Amplified	168 (24%)
Unknown	7
DNA Index	
Hyperdiploid	393 (63%)
Diploid	227 (37%)
Unknown	74
Histology	
Favorable	224 (36%)
Unfavorable	392 (64%)
Unknown	78
Risk	
Low/Intermediate	221 (32%)
High	470 (68%)
Unknown	3

4.2 Copy number alteration detection in primary tumors

Since measuring copy number (CN) as a discrete variable does not account for tumor heterogeneity, we used the OverUnder algorithm, a previously described copy number alteration (CNA) detection algorithm [16] for Illumina SNP arrays, for determining somatically acquired CNAs in primary tumors from neuroblastoma patients. Briefly, after correcting probe intensity values for aneuploidy, absolute copy number was determined for each 101-SNP window based on its allelic ratio. Adjacent SNPs were grouped into regions. Regions whose absolute copy number was more than 0.8 greater than the baseline copy number (based on the cell's computed DNA index) were called gains. The absolute copy number for each region is then normalized against the ploidy measure for each sample to obtain the relative copy number estimates. The performance of the algorithm has been previously validated using primary tumor from neuroblastoma patients with qPCR-inferred copy number estimates [16].

4.3 GISTIC analysis of recurrent somatic copy number alterations

We assessed the significance of recurrent genomic alterations using the GISTIC algorithm [17] (version 3) which is part of the Gene Pattern Server at the Broad Institute (http://genepattern.broadinstitute.org). Relative copy number estimates for 474 high-risk neuroblastoma tumors were generated using the OverUnder algorithm in an unmatched (to blood) analysis in order to optimize the number of usable samples. To exclude constitutional variants from our tumor copy number analysis, we constructed a CNV filter file for GISTIC that included all CNVs identified with a frequency greater than one percent in our recent CNV-GWAS [18] (file available upon request).

5 Survival analysis on SNP genotype and copy number alteration

5.1 Patient and tumor samples

The clinical data for survival analysis is retrieved from the Children's Oncology Group (COG), so the analysis is restricted to samples with available data. More specifically, out of the 1,597 cases with whole-genome genotype data, 68 did not have survival information and were eliminated, leaving 1,529 used for the survival analysis on SNP genotypes. Out of these 68 patients, 12 were never enrolled on a COG study, three only enrolled on a non-therapeutic study and 53 were pre-registered patients who never fully enrolled. For the survival analysis on tumor DNA copy number, among 701 patients with copy number estimates, seven patients were eliminated because they had no outcome data and limited risk factors. This left 694 patients available for analysis.

5.2 Statistical approach

For event-free survival (EFS), time to event was defined as the time from diagnosis until the time of first occurrence of relapse, progressive disease, secondary malignancy, or death, or until the time of last contact if no event occurred. Patients who were alive without event were censored at the time last known alive. For overall survival (OS), death was the only event considered. Survival analyses were performed using the methods of Kaplan and Meier, with standard errors per the methods of Peto et al. [19].

6 Microarray expression data analysis

6.1 Data source

We quantified mRNA expression in a highly annotated series of 101 prospectively collected diagnostic neuroblastoma primary tumors and the expression profiles were determined using Affymetrix U95Av2 arrays [20]. These primary tumor samples were selected from the COG (n = 91 prospectively collected) or Children's Hospital of Philadelphia (n = 10) neuroblastoma tumor banks. Samples were selected so that a minimum of 20 cases would be available in four clinically and biologically distinct subsets of neuroblastoma: low-risk, intermediate-risk, high-risk, and high-risk with *MYCN* amplification. This data set was previously used to demonstrate that the genomic data can be used to subcategorize the disease into molecular subsets and the regional copy number alterations are correlated with a broad number of transcriptional alterations genome wide [20]. The raw data set as CEL files were available from the GEO database with accession number GSE3960. A small subset (n=61) of these tumor samples have paired DNA samples from whole blood and also paired DNA samples from tumor, which were used in the genotype-expression association analysis.

6.2 Genotype-expression association

The expression measures for each probe set in the Affymetrix array is extracted and normalized using well-established Robust Multi-array Average (RMA) protocols[21] from raw CEL files. The latest probe set annotation file (revision na27) from Affymetrix's website (http://www.affymetrix.com) was used to assign expression value for each probe set to the corresponding gene. The genotype-expression association for SNPs was performed using linear regression, and the association for copy number was performed using two-sided t-test.

7 RT-PCR to measure gene expression levels

Real-time quantitative PCR (RT-PCR) was used for measuring gene expression levels in a subset of tumor samples without *LMO1* gain, to assess the relationships between *LMO1* risk genotypes and expression levels. All primer/probe sets spanned exon boundaries to assure specificity for cDNA. The probes were purchased from Applied Biosystems (Foster City, CA) as on-the-shelf product with assay ID **Hs00231133_m1**. Relative expression of target gene was determined by normalization to glyceraldehyde-3-phosphate dehydrogenase (GAPDH) or Hypoxanthine-guanine phosphoribosyltransferase (HPRT) using a standard curve method with 10 serial dilutions according to the manufacturer's instruction. All RT-PCR experiments included a no template control and were done in triplicate.

8 *LMO1* knockdown and over-expression

For *LMO1* knockdown, cells were infected with control copGFP lentiviral particles (Santa Cruz,cat # sc-108084) or *LMO1* shRNA(h) lentiviral particles (Santa Cruz,cat # sc-38025-v). On day 0, cell lines were seeded on a 12 well plate in RPMI 1640 complete

media (Gibco 22400) containing 10% fetal bovine serum (Hyclone SH 30073-03), 1X antibiotic antimycotic (Gibco 15240-062), 2 mM L-Glutamine (Cell Grow 25-005-CI) and 50ug/ml gentamycin (Gibco 15750-060) such that the cells were 50% confluent on day 1. On day 1, media was removed and replaced with 1.0 ml RPMI 1640 complete media containing 5ug/ml polybrene (sc-134220). Next 20 ul of lentiviral particles were added to the media with gentle swirling and incubated overnight. On day 3, media was replaced with fresh RPMI 1640 complete media. On day 5, infected cells were selected by passaging into 6 well plates in RPMI 1640 complete media containing 1ug/ml puromycin (Sigma P9620). Infection efficiency was monitored by GFP fluorescence and was nearly 90%. These surviving cells were pooled for experiments. Real-time cell growth was monitored once every hour for 96 consecutive hours using the RT-CES system (ACEA Biosciences), as previously described [22-24]. Six replicates of each treatment were plated at cell densities of 22,500 per well of a 96-well plate for LAN5, 12,500 for SKNSH, and 10,000 for BE-2C and NLF. Cell growth curves were normalized at 8hr post seeding. Markers represent the average normalized cell index ± standard error.

For overexpression of LMO1, we used pcDNA3-LMO1 vector previously described, with pcDNA3-vector as a control [25]. On day 0, $3x10^6$ SK-N-BE2C neuroblastoma cells were seeded in a 100 mm dish. On day 1, plasmids were transfected using a 6:1 ratio of Fugene 6:plasmid as described in the Fugene protocol. Briefly, 36 ul of Fugene was added to 564 ul of RPMI 1640 serum-free media and incubated for 5 minutes. Six ug of plasmid DNA was then added and the resulting complex was incubated for 15 minutes at room temperature and added to cells. On day 3, 400 ug/mL G418 (Cellgro Cat# 30-234-CR) was added to cells, which were selected for 4 weeks. All colonies were pooled for subsequent experiments. Real-time cell growth was monitored once every hour for 96 consecutive hours using the RT-CES system (ACEA Biosciences). Twelve replicates of

each treatment were plated at a cell density of 10,000 cells per well of a 96-well plate in RPMI 1640 complete media. Growth curves were normalized at 8hrs post seeding. Markers represent the average normalized cell index ± standard error.

9 Supplementary Tables

Supplementary Table 1. Association results at the previously reported 6p22 locus for SNPs that reaches genome-wide significance.

Discovery Cohort								US Replication Cohort				Combined	
SNP	Position	A1/A2[1]	Function	A1 Freq in cases	A1 Freq in controls	Allelic P-value	OR	A1 Freq in cases	A1 Freq in controls	Allelic P-value	OR	CMH P-value[2]	CMH OR[3]
rs4712653	22233943	C/T	intron	0.5418	0.4593	2.46E-14	1.392	0.5447	0.4526	0.000684	1.447	7.53E-17	1.399
rs9295536	22239908	A/C	intron	0.5111	0.4331	3.25E-13	1.368	0.5158	0.418	0.000286	1.483	5.29E-16	1.383

[1]: A1: Allele 1; A2: Allele 2.

[2]: CMH: Cochran-Mantel-Haenszel test.

[3]: OR: odds ratio (cases versus controls for A1)

Supplementary Table 2. Association results at the previously reported 2q35 locus for SNPs that reaches genome-wide significance.

Discovery Cohort								US Replication Cohort				Combined	
SNP	Position	A1/A2[1]	Function	A1 Freq in cases	A1 Freq in controls	Allelic P-value	OR	A1 Freq in cases	A1 Freq in controls	Allelic P-value	OR	CMH P-value[2]	CMH OR[3]
rs3768716	215344039	C/T	intron	0.2938	0.2272	7.25E-13	1.416	0.2526	0.2196	0.1456	1.201	5.07E-13	1.386
rs17487792	215351745	T/C	intron	0.292	0.2262	1.35E-12	1.411	0.2526	0.2174	0.1193	1.217	6.96E-13	1.385
rs7587476	215362132	T/C	intron	0.3162	0.2466	3.05E-13	1.413	0.2751	0.2455	0.2093	1.166	3.96E-13	1.378
rs6712055	215375149	C/T	intron	0.3629	0.3023	2.17E-09	1.314	0.3605	0.2952	0.008939	1.346	6.79E-11	1.319
rs6435862	215380791	G/T	intron	0.3456	0.2869	2.93E-09	1.313	0.3368	0.2631	0.002317	1.423	3.35E-11	1.327
rs6715570	215381685	T/C	intron	0.3663	0.3052	1.21E-09	1.316	0.3605	0.2864	0.002835	1.405	1.54E-11	1.328
rs2592232	215433996	G/A	Intergenic	0.371	0.309	8.87E-10	1.319	0.3605	0.3109	0.0499	1.25	1.35E-10	1.31
rs10498025	215457501	G/A	Intergenic	0.3165	0.2537	5.45E-11	1.362	0.2974	0.2502	0.04678	1.269	8.26E-12	1.35
rs10498026	215457595	A/G	Intergenic	0.4142	0.4756	9.69E-09	0.7798	0.4316	0.4801	0.07431	0.8222	2.10E-09	0.7854

[1]: A1: Allele 1; A2: Allele 2.

[2]: CMH: Cochran-Mantel-Haenszel test.

[3]: OR: odds ratio (cases versus controls for A1)

Supplementary Table 3. Pairwise interaction tests between the most significant markers at four loci (3 SNPs, 1 CNV) using case/control data failed to identify any significant interaction. Only 1,420 cases and 3,082 control subjects with CNV information in the Discovery case series were used in the association analysis. OR_INT represents odds ratio, with 1 indicating lack of epistasis. PLINK was used for the calculation of P-values.

CHR1	SNP1	CHR2	SNP2	OR_INT	P
1	1q21.1cnv	2	rs7587476	0.9259	0.5146
1	1q21.1cnv	6	rs4712653	1.174	0.1367
1	1q21.1cnv	11	rs110419	0.9374	0.5344
2	rs7587476	6	rs4712653	1.018	0.7936
2	rs7587476	11	rs110419	1.016	0.8166
6	rs4712653	11	rs110419	0.923	0.1957

Supplementary Table 4. Pairwise interaction tests between four markers using case-only data failed to identify any significant interaction. Only 1,420 cases and 3,082 control subjects with CNV information in the Discovery case series were used in the association analysis. PLINK was used for the calculation of P-values.

CHR1	SNP1	CHR2	SNP2	P
1	1q21.1cnv	2	rs7587476	0.07821
1	1q21.1cnv	6	rs4712653	0.8632
1	1q21.1cnv	11	rs110419	0.3129
2	rs7587476	6	rs4712653	0.1051
2	rs7587476	11	rs110419	0.735
6	rs4712653	11	rs110419	0.3526

Supplementary Table 5. A list of genotyped and imputed SNP markers on 11p15.4 that have P-values less than $1x10^{-4}$ in the combined association analysis on US discovery and replication case series with whole-genome genotype data. Odds ratios (OR) were calculated with respect to allele 1 (A1).

SNP	Position	P-value	Type	A1	A2	A1 freq in cases	A1 freq in controls	OR
rs110420	8209625	1.17E-13	imputed	C	T	0.4395	0.5123	0.7422
rs110419	8209429	1.35E-13	typed	G	A	0.4395	0.5123	0.7422
rs4758317	8207387	1.74E-13	imputed	C	A	0.5498	0.4744	1.356
rs204928	8211009	2.41E-13	imputed	G	A	0.4352	0.5076	0.7427
rs204926	8211682	6.25E-12	imputed	A	G	0.4372	0.5054	0.7543
rs4758051	8195215	6.94E-11	typed	G	A	0.5146	0.4504	1.298
rs4758050	8195121	1.25E-10	imputed	G	C	0.5133	0.4492	1.297
rs10840000	8196689	1.89E-10	imputed	G	C	0.5155	0.4513	1.297
rs417210	8225981	7.22E-08	imputed	G	T	0.3744	0.3198	1.265
rs10840002	8199602	4.03E-07	typed	A	G	0.4146	0.364	1.24
rs204938	8234773	1.79E-06	typed	C	T	0.4894	0.4455	1.201
rs11041820	8208014	6.80E-06	typed	A	G	0.3046	0.2609	1.227
rs204937	8237514	3.07E-05	typed	C	T	0.46	0.4163	1.201
rs3794012	8226820	3.22E-05	typed	C	T	0.3836	0.4303	0.8147
rs11601177	8188761	4.36E-05	imputed	C	G	0.2364	0.2031	1.227
rs12362235	8233234	6.27E-05	imputed	G	A	0.3916	0.4313	0.845
rs7951027	8190452	6.72E-05	typed	A	G	0.3267	0.289	1.18

Supplementary Table 6. Correlation of rs110419 genotypes (AA, AG, GG) with clinical variables.

				P-value [2]	
	AA[1]	**AG[1]**	**GG[1]**	**AA vs GG**	**AG vs GG**
Stage 4	235 (35%)	323 (48%)	117 (17%)		
Not Stage 4	247 (30%)	390 (47%)	191 (23%)	0.0030	0.0303
***MYCN* Amp**	74 (28%)	129 (50%)	58 (22%)		
***MYCN* Not Amp**	381 (33%)	539 (47%)	233 (20%)	0.2004	0.8237
High risk	232 (35%)	317 (48%)	110 (17%)		
Not High risk	241 (30%)	385 (47%)	189 (23%)	0.0008	0.0143
DNA Index Hyperdiploid	296 (34%)	401 (45%)	187 (21%)	0.8401	0.1586
DNA Index Diploid	144 (30%)	249 (51%)	94 (19%)		
Unfavorable Histology	179 (33%)	260 (48%)	104 (19%)	0.2277	0.1426
Favorable Histology	218 (32%)	308 (45%)	154 (23%)		
Age >= 1 year	336 (34%)	477 (49%)	168 (17%)	<0.0001	0.0006
Age < 1 year	155 (28%)	251 (46%)	142 (26%)		

1: The AA, AG and GG genotypes represent homozygous risk genotype, heterozygous risk genotype and non-risk genotype, respectively.
2: two-sided Fisher's exact test

Supplementary Table 7. Correlation of rs110419 risk allele with clinical variables

	A[1]	G[1]	P-value[2]	OR (A vs G)
Stage 4	793 (59%)	557 (41%)	0.0040	1.24
Not Stage 4	884 (53%)	772 (47%)		
***MYCN* Amp**	277 (53%)	245 (47%)	0.1720	0.87
***MYCN* Not Amp**	1301 (56%)	1005 (44%)		
High risk	781 (59%)	537 (41%)	0.0010	1.28
Not High risk	867 (53%)	763 (47%)		
DNA Index Hyperdiploid	993 (56%)	775 (44%)	0.6297	1.04
DNA Index Diploid	537 (55%)	437 (45%)		
Unfavorable Histology	618 (57%)	468 (43%)	0.2870	1.10
Favorable Histology	744 (55%)	616 (45%)		
Age >= 1 year	1149 (59%)	813 (41%)	<0.0001	1.35
Age < 1 year	561 (51%)	535 (49%)		

1: A and G represent risk and non-risk allele for rs110419, respectively.
2: two-sided Fisher's exact test

Supplementary Table 8. Survival analysis stratified by rs110419 Genotypes

Patient cohort	n (%)	5-year EFS +/- std error	EFS P-value [1]	5-year OS +/- std error	OS p-value [2]
Overall	1529	67 ± 2	N/A	73 ± 2	N/A
rs110419					
GG	310 (20%)	75 ± 3	0.0085	81 ± 3	0.0217
AG	728 (48%)	66 ± 3		71 ± 2	
AA	491 (32%)	64 ± 3		71 ± 3	

1. Event-free survival logrank test
2. Overall survival logrank test

Supplementary Table 9. Correlation of *LMO1* copy number status in diagnostic tumor materials with clinical and biological covariates.

Biology Variable	No Segmental Gain	Segmental Gain	p-value[1]
Age < year	161	7	0.0001
Age ≥ 1 year	448	78	
Stage 1,2,3,4s	239	8	< 0.0001
Stage 4	369	76	
MYCN Not Amplified	440	79	< 0.0001
MYCN Amplified	165	3	
Hyperdiploid	348	45	0.6896
Diploid	204	23	
Histology Favorable	211	13	0.0013
Histology Unfavorable	336	56	
Low/Intermediate-Risk	212	9	< 0.0001
High-Risk	395	75	

1: P-values were calculated by 2-sided Fisher's exact test.

Supplementary Table 10. Survival analysis stratified by *LMO1* copy number in tumors.

Patient cohort	n (%)	5-year EFS +/- std error	EFS p-value[1]	5-year OS +/- std error	OS p-value[2]
Overall	694	52 ± 3	N/A	63 ± 2	N/A
LMO1 No Segmental Gain Segmental Gain	609 (88%) 85 (12%)	53 ± 3 40 ± 7	0.0765	65 ± 3 52 ± 7	0.0411

1. Event-free survival logrank test
2. Overall survival logrank test

Supplementary Table 11. Gene expression and Western blot of *LMO1* in neuroblastoma cell lines with homozygous genotypes. The *LMO1* expression is measured by TaqMan and is normalized against *HPRT1*; the densitometry for western blot is normalized against actin.

rs110419 Illumina genotype	**Sample Name (genotype)**	***LMO1* mRNA expression**	***LMO1* Western**
AA	LAN-5 (AA)	1.478095	6.505809
AA	KCN (AA)	1.264893	5.826934
AA	SKNSH (AA)	1.307877	4.497628
AA	EBC-1 (AA)	1.293404	1.54432
AA	SKNAS (AA)	1.001759	1.454939
BB	CHP 134 (GG)	0.682634	0.46277
BB	IMR-5 (GG)	0.488344	0.039363
BB	BE2C (GG)	0.896645	0.269748
BB	NLF (GG)	0.791417	0.013708

Supplementary Table 12. Raw gene expression values of *LMO1* in 61 tumors, who have been genotyped by Illumina SNP arrays, whose gene expression is profiled by Affymetrix U95Av2 microarrays, and whose paired blood samples have been genotyped by Illumina SNP arrays. A second probe set (35104_r_at) is spotted at the Affymetrix array, but it is a "rules dropped" probe and does not measure *LMO1* expression accurately.

Sample	Tumor	LMO1 duplication	rs110419 genotype	35103_i_at
GSM90306	37	0	1	8.21336
GSM90308	406	0	1	7.85906
GSM90309	491	0	1	7.79781
GSM90316	1073	0	1	7.25909
GSM90324	1256	0	1	5.7843
GSM90325	1260	0	1	7.83856
GSM90326	1285	0	1	7.72689
GSM90339	417	0	1	7.3324
GSM90344	1133	0	1	7.78378
GSM90346	1196	0	1	7.05523
GSM90351	1292	0	1	7.31254
GSM90357	158	0	1	8.36897
GSM90359	338	0	1	7.56287
GSM90362	495	0	1	6.78126
GSM90369	1100	0	1	6.72926
GSM90375	1204	0	1	7.90693
GSM90380	1258	0	1	6.58286
GSM90381	1310	0	1	7.58696
GSM90389	433	0	1	7.47246
GSM90395	1066	0	1	8.28722
GSM90396	1068	0	1	5.61129
GSM90398	1129	0	1	7.45884
GSM90400	1246	0	1	4.66561
GSM90310	989	0	2	6.25926
GSM90312	1013	0	2	7.85804
GSM90313	1030	0	2	8.04233
GSM90314	1033	0	2	7.20403
GSM90335	15	0	2	7.26725
GSM90347	1250	0	2	4.63741
GSM90349	1290	0	2	7.48474
GSM90352	1303	0	2	8.88454
GSM90360	396	0	2	6.78135
GSM90371	1156	0	2	6.19723
GSM90388	260	0	2	6.12032
GSM90391	969	0	2	6.07663

GSM90402	1266	0	2	4.92303
GSM90405	1749	0	2	5.13585
GSM90315	1059	0	3	6.02265
GSM90323	1254	0	3	7.63698
GSM90327	1319	0	3	6.72812
GSM90328	1322	0	3	6.06761
GSM90330	1556	0	3	6.84404
GSM90332	2175	0	3	7.67793
GSM90336	46	0	3	4.81441
GSM90337	975	0	3	7.49569
GSM90338	415	0	3	4.76739
GSM90361	427	0	3	8.1236
GSM90373	1174	0	3	5.90417
GSM90345	1163	1	1	7.98167
GSM90355	58	1	1	8.72906
GSM90364	974	1	1	7.4473
GSM90372	1167	1	1	7.16587
GSM90374	1194	1	1	8.08574
GSM90376	1213	1	1	8.57743
GSM90377	1218	1	1	7.24941
GSM90365	982	1	2	7.26181
GSM90367	1092	1	2	8.07874
GSM90368	1097	1	2	7.65596
GSM90378	1220	1	2	8.06886
GSM90318	1139	1	3	5.73713
GSM90340	430	1	3	8.30058

Supplementary Table 13. Sanger sequencing of the 1kb region at *LMO1* promoter on 20 cell lines (2 DNA samples failed), including 10 with homozygous risk alleles and 10 with homozygous non-risk alleles. It is likely that genetic variants at other regulatory regions are responsible for the association signal.

Variant	Type	Observed counts
-380insA	Insertion	18
-379insG	Insertion	18
-479T->G	Single nucleotide variant	2
-110G->T	Single nucleotide variant	1
-380insA, -379insG	Insertion	10 in 10 additional randomly chosen blood samples

Supplementary Table 14. SNPs that are in moderate LD (D'>0.5) with rs110419, based on the 1000 Genomes Project data on 120 phased haplotypes from CEU subjects (2010 release). These SNPs were located between chr11:8100088 and chr11:8341655 (*LMO1* plus 100kb upstream/downstream region). Two synonymous SNPs (rs1042359 and rs3750952) were located within *LMO1*, while non-synonymous SNPs were not identified.

SNP	MAF	r^2	D'
chr11:8100088	0.07	0.03	0.73
chr11:8102026	0.06	0.02	0.69
chr11:8102633	0.06	0.02	0.69
rs11041753	0.02	0.01	1.00
rs11041755	0.02	0.01	1.00
rs3750954	0.08	0.04	0.76
rs1528133	0.06	0.02	0.69
chr11:8108746	0.02	0.01	1.00
chr11:8109391	0.06	0.02	0.69
chr11:8110543	0.06	0.02	0.69
rs11041758	0.02	0.01	1.00
chr11:8110895	0.02	0.02	1.00
rs11041759	0.08	0.04	0.76
rs1528128	0.06	0.02	0.69
chr11:8113938	0.01	0.01	1.00
chr11:8114662	0.06	0.02	0.69
rs11041761	0.02	0.01	1.00
rs11041762	0.02	0.01	1.00
chr11:8116193	0.06	0.02	0.69
chr11:8116433	0.06	0.02	0.69
chr11:8116460	0.03	0.02	1.00
chr11:8116537	0.06	0.02	0.69
rs3812761	0.05	0.04	1.00
rs6578933	0.08	0.04	0.76
rs9783349	0.08	0.04	0.76
chr11:8124642	0.02	0.01	1.00
chr11:8126270	0.05	0.04	1.00
chr11:8126464	0.06	0.02	0.69
rs12098876	0.06	0.02	0.69
chr11:8126788	0.02	0.02	1.00
chr11:8128145	0.06	0.02	0.69
rs12421498	0.11	0.11	0.86
rs11820921	0.06	0.02	0.69
rs10839983	0.08	0.04	0.76
rs11041770	0.02	0.01	1.00
chr11:8131874	0.06	0.02	0.69
chr11:8132038	0.06	0.02	0.69

chr11:8132070	0.06	0.02	0.69
rs16936374	0.06	0.02	0.69
chr11:8134871	0.05	0.04	1.00
chr11:8135154	0.06	0.02	0.69
chr11:8136055	0.06	0.02	0.69
chr11:8137349	0.06	0.02	0.69
rs4758298	0.06	0.02	0.69
chr11:8139318	0.03	0.03	1.00
rs4758299	0.06	0.02	0.69
chr11:8139590	0.11	0.11	0.86
rs16936453	0.06	0.02	0.69
chr11:8140267	0.03	0.03	1.00
chr11:8140283	0.06	0.02	0.69
chr11:8140348	0.06	0.02	0.69
chr11:8140565	0.05	0.04	1.00
chr11:8140689	0.05	0.04	1.00
rs16936464	0.05	0.04	1.00
chr11:8141376	0.05	0.04	1.00
chr11:8145275	0.06	0.02	0.69
rs11825426	0.05	0.04	1.00
rs11825474	0.06	0.02	0.69
chr11:8146489	0.11	0.11	0.86
chr11:8146750	0.05	0.04	1.00
rs7113581	0.05	0.04	1.00
rs4758301	0.05	0.04	1.00
rs4758045	0.03	0.02	1.00
rs7942987	0.02	0.01	1.00
rs10769880	0.05	0.04	1.00
chr11:8150015	0.06	0.05	1.00
chr11:8150147	0.06	0.05	1.00
rs6578937	0.03	0.02	1.00
rs11041782	0.14	0.08	0.74
rs7937200	0.02	0.01	1.00
rs7930523	0.01	0.01	1.00
chr11:8152523	0.08	0.03	0.59
chr11:8154726	0.05	0.04	1.00
chr11:8156573	0.08	0.02	0.52
chr11:8156892	0.03	0.02	1.00
chr11:8158039	0.02	0.02	1.00
chr11:8159626	0.04	0.01	0.56
rs12365076	0.10	0.06	0.69
rs3849987	0.08	0.02	0.52
chr11:8169127	0.03	0.01	0.54

chr11:8170696	0.14	0.05	0.61
chr11:8170697	0.14	0.05	0.61
rs16936813	0.15	0.06	0.64
chr11:8172806	0.16	0.07	0.66
chr11:8173224	0.17	0.08	0.67
chr11:8173742	0.17	0.08	0.67
chr11:8176434	0.08	0.03	0.59
rs3893363	0.17	0.05	0.56
chr11:8181651	0.17	0.05	0.56
rs10839998	0.16	0.06	0.51
rs1881229	0.16	0.08	0.61
chr11:8186928	0.05	0.04	1.00
chr11:8187976	0.22	0.16	0.83
rs11041808	0.48	0.27	0.54
rs11041809	0.46	0.30	0.64
rs11041810	0.43	0.27	0.66
rs11041811	0.44	0.30	0.67
rs11041812	0.44	0.30	0.67
chr11:8188296	0.04	0.04	1.00
rs11607164	0.26	0.22	0.86
rs11607616	0.26	0.22	0.86
rs11601177	0.24	0.19	0.85
rs4758313	0.02	0.01	1.00
rs4758314	0.28	0.24	0.87
chr11:8188954	0.23	0.17	0.84
chr11:8189751	0.04	0.01	0.56
rs7483500	0.02	0.01	1.00
chr11:8190430	0.03	0.03	1.00
rs7951027	0.33	0.28	0.83
chr11:8190745	0.25	0.21	0.85
chr11:8191111	0.03	0.03	1.00
rs11041813	0.43	0.31	0.70
rs7927509	0.03	0.03	1.00
rs10839999	0.43	0.32	0.71
rs10769885	0.45	0.32	0.68
rs34135547	0.10	0.06	0.82
rs12806081	0.08	0.05	0.78
chr11:8193649	0.08	0.05	0.78
rs4758049	0.45	0.32	0.68
rs4758050	0.46	0.30	0.64
rs4758051	0.45	0.32	0.68
rs7929853	0.24	0.19	0.85
chr11:8195719	0.02	0.02	1.00

chr11:8195815	0.03	0.03	1.00
chr11:8195914	0.02	0.01	1.00
rs7122225	0.24	0.19	0.85
chr11:8196575	0.08	0.05	0.78
rs10840000	0.46	0.30	0.64
rs7933766	0.47	0.28	0.61
chr11:8197049	0.08	0.05	0.78
chr11:8198408	0.07	0.03	0.73
chr11:8198782	0.17	0.07	0.54
chr11:8198798	0.13	0.04	0.51
rs3861073	0.09	0.03	0.60
rs7109806	0.13	0.12	1.00
rs7106487	0.14	0.14	1.00
rs10840002	0.32	0.14	0.60
rs35063107	0.02	0.01	1.00
rs6578942	0.13	0.12	1.00
rs767838	0.13	0.13	1.00
rs11041816	0.48	0.44	0.69
rs1042359	0.23	0.21	0.92
rs3816490	0.09	0.09	1.00
rs12795672	0.06	0.07	1.00
rs12283881	0.09	0.09	1.00
rs4315061	0.44	0.93	1.00
rs1037514	0.24	0.27	1.00
rs7933956	0.13	0.13	1.00
rs11041818	0.03	0.03	1.00
rs2290451	0.31	0.33	0.94
rs2290450	0.31	0.33	0.94
rs7482131	0.13	0.13	1.00
rs4758053	0.13	0.13	1.00
chr11:8206548	0.03	0.03	1.00
rs7952320	0.46	0.93	0.97
rs34483450	0.07	0.05	0.77
chr11:8207227	0.13	0.13	1.00
rs12278655	0.11	0.10	1.00
rs4758317	0.46	0.93	0.97
rs9943637	0.13	0.13	1.00
rs11041820	0.31	0.33	0.94
rs10734630	0.13	0.13	1.00
rs3750952	0.45	0.97	1.00
chr11:8209080	0.05	0.04	1.00
rs12577949	0.05	0.04	1.00
chr11:8209508	0.06	0.05	1.00

rs110420	0.45	0.97	1.00
rs16937174	0.13	0.13	1.00
rs204928	0.43	0.90	1.00
rs12574271	0.32	0.31	0.89
rs204927	0.11	0.10	1.00
rs12577773	0.32	0.31	0.89
rs204926	0.44	0.93	1.00
rs2168101	0.28	0.47	1.00
rs7944727	0.43	0.54	0.92
rs7948497	0.50	0.73	0.93
chr11:8214346	0.02	0.01	1.00
rs7121476	0.36	0.20	0.64
rs12269893	0.04	0.01	0.56
chr11:8215140	0.03	0.04	1.00
rs439667	0.28	0.21	0.80
rs411655	0.04	0.01	0.56
rs1794087	0.02	0.02	1.00
chr11:8216881	0.08	0.07	0.82
rs12295662	0.06	0.07	1.00
rs413708	0.10	0.06	0.69
chr11:8217911	0.02	0.01	1.00
rs12277194	0.06	0.07	1.00
chr11:8218849	0.06	0.05	1.00
rs3849994	0.10	0.06	0.82
rs17227935	0.32	0.23	0.77
chr11:8219594	0.29	0.23	0.81
chr11:8219909	0.07	0.02	0.54
chr11:8219924	0.07	0.03	0.73
chr11:8220165	0.03	0.02	1.00
chr11:8220504	0.05	0.03	0.69
rs442264	0.41	0.24	0.55
rs484161	0.38	0.32	0.67
rs3794015	0.47	0.28	0.61
rs376813	0.04	0.04	1.00
rs12278467	0.16	0.05	0.54
rs12277079	0.16	0.05	0.54
chr11:8222244	0.28	0.22	0.81
rs3794014	0.41	0.32	0.62
rs2061036	0.10	0.06	0.82
rs2061037	0.50	0.24	0.53
rs2061038	0.10	0.06	0.82
chr11:8223570	0.02	0.01	1.00
rs10840003	0.50	0.30	0.60

rs444284	0.12	0.03	0.53
rs452348	0.10	0.06	0.82
rs391561	0.33	0.30	0.84
rs7940839	0.06	0.07	1.00
rs404989	0.33	0.30	0.84
rs417210	0.33	0.30	0.84
chr11:8226110	0.02	0.02	1.00
chr11:8226259	0.02	0.02	1.00
rs11041823	0.06	0.07	1.00
rs450319	0.28	0.21	0.80
chr11:8226667	0.02	0.02	1.00
rs3794012	0.43	0.28	0.57
rs11041824	0.06	0.07	1.00
rs455082	0.10	0.06	0.82
rs34494077	0.28	0.24	0.87
chr11:8227543	0.02	0.01	1.00
rs11604384	0.10	0.06	0.82
rs11041825	0.10	0.06	0.82
chr11:8228510	0.06	0.07	1.00
chr11:8228741	0.05	0.04	1.00
rs379951	0.48	0.29	0.62
rs6578947	0.10	0.06	0.82
rs7935649	0.10	0.06	0.82
chr11:8230041	0.12	0.08	0.74
chr11:8230042	0.12	0.08	0.74
chr11:8230045	0.10	0.06	0.69
chr11:8230116	0.27	0.23	0.86
chr11:8230258	0.02	0.02	1.00
rs1598165	0.27	0.23	0.86
chr11:8230976	0.02	0.02	1.00
chr11:8231010	0.05	0.04	1.00
rs11606296	0.10	0.06	0.82
rs4237769	0.43	0.26	0.54
chr11:8232365	0.02	0.02	1.00
rs1454438	0.07	0.05	0.77
rs12362235	0.43	0.26	0.54
rs7106955	0.13	0.05	0.54
chr11:8233292	0.28	0.24	0.87
rs11041829	0.43	0.26	0.54
rs7111233	0.18	0.06	0.58
chr11:8234238	0.16	0.09	0.77
chr11:8234549	0.05	0.04	1.00
chr11:8234652	0.02	0.01	1.00

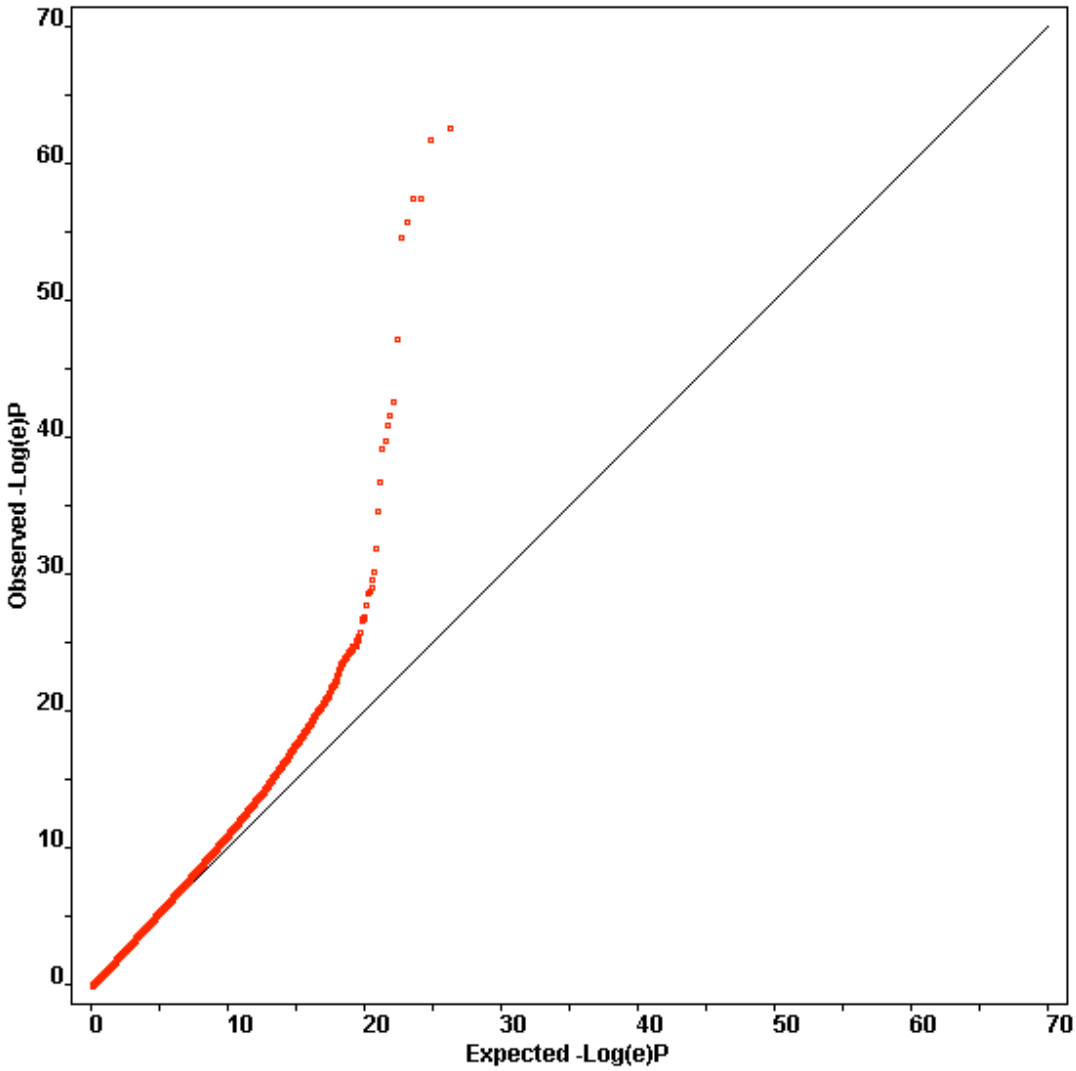
rs204938	0.48	0.43	0.74
rs11605629	0.11	0.05	0.66
rs204937	0.48	0.37	0.69
chr11:8237557	0.06	0.07	1.00
rs3794010	0.31	0.22	0.76
chr11:8238216	0.05	0.04	1.00
chr11:8238608	0.05	0.04	1.00
chr11:8239287	0.36	0.14	0.54
chr11:8239914	0.29	0.14	0.63
chr11:8243864	0.12	0.06	0.60
chr11:8244455	0.03	0.02	1.00
chr11:8247776	0.03	0.04	1.00
chr11:8249593	0.03	0.04	1.00
rs11822564	0.03	0.04	1.00
rs7931117	0.03	0.03	1.00
rs7128430	0.08	0.07	1.00
rs7105921	0.08	0.07	1.00
rs4758318	0.03	0.04	1.00
rs4758319	0.03	0.04	1.00
rs11041841	0.07	0.06	1.00
chr11:8256808	0.03	0.03	1.00
chr11:8256850	0.02	0.01	1.00
rs7111412	0.03	0.04	1.00
chr11:8257236	0.03	0.04	1.00
chr11:8257697	0.03	0.04	1.00
rs7104520	0.10	0.06	0.82
chr11:8257998	0.03	0.04	1.00
chr11:8258331	0.03	0.04	1.00
chr11:8264528	0.07	0.03	0.73
rs7952235	0.06	0.05	1.00
rs35217076	0.03	0.01	0.54
chr11:8273489	0.05	0.02	0.64
chr11:8273508	0.01	0.01	1.00
rs11041852	0.06	0.05	1.00
rs12291703	0.06	0.05	1.00
chr11:8278333	0.08	0.04	0.63
rs7112171	0.06	0.05	1.00
rs9804641	0.06	0.05	1.00
chr11:8279065	0.10	0.04	0.54
rs9804435	0.06	0.05	1.00
rs12222419	0.02	0.02	1.00
chr11:8283372	0.06	0.05	1.00
rs10840014	0.06	0.05	1.00

chr11:8288766	0.03	0.04	1.00
rs34099220	0.05	0.03	0.69
rs16937623	0.02	0.02	1.00
rs10743069	0.03	0.02	1.00
chr11:8292477	0.03	0.02	1.00
chr11:8292482	0.03	0.04	1.00
chr11:8294707	0.03	0.02	1.00
chr11:8296367	0.10	0.04	0.64
chr11:8303239	0.07	0.02	0.54
chr11:8308133	0.05	0.02	0.64
rs35458360	0.13	0.04	0.51
rs4485100	0.13	0.04	0.51
chr11:8312049	0.11	0.05	0.57
chr11:8315762	0.06	0.04	0.74
rs36087766	0.09	0.05	0.66
rs35126723	0.08	0.06	0.79
chr11:8320296	0.03	0.02	1.00
rs4272776	0.08	0.06	0.79
rs4409808	0.08	0.06	0.79
rs4465360	0.08	0.06	0.79
rs4463834	0.08	0.06	0.79
rs4468340	0.08	0.06	0.79
chr11:8324919	0.02	0.02	1.00
rs4462333	0.08	0.06	0.79
rs4258369	0.08	0.06	0.79
rs4474423	0.08	0.06	0.79
rs4316500	0.08	0.06	0.79
rs4414207	0.08	0.06	0.79
rs6578961	0.08	0.06	0.79
rs7930847	0.08	0.06	0.79
rs12802666	0.08	0.06	0.79
rs35558051	0.08	0.06	0.79
rs34161135	0.08	0.06	0.79
rs12785737	0.09	0.05	0.66
rs12808792	0.09	0.05	0.66
rs11606763	0.08	0.06	0.79
rs11605336	0.08	0.06	0.79
chr11:8333020	0.08	0.06	0.79
chr11:8334536	0.08	0.06	0.79
rs4279991	0.08	0.06	0.79
rs35717935	0.08	0.06	0.79
rs35932155	0.07	0.05	0.77
rs34843885	0.08	0.06	0.79

rs4441008	0.08	0.06	0.79
rs4316501	0.08	0.06	0.79
rs36104415	0.05	0.03	0.69
chr11:8340149	0.03	0.02	1.00
chr11:8340834	0.07	0.05	0.77

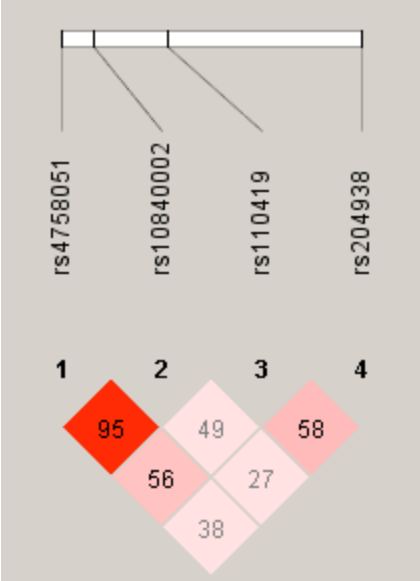
10 Supplementary Figures

Supplementary Figure 1. Quantile-Quantile plots on the expected and observed P-values for SNPs passing quality control. The genomic inflation factor is 1.08.

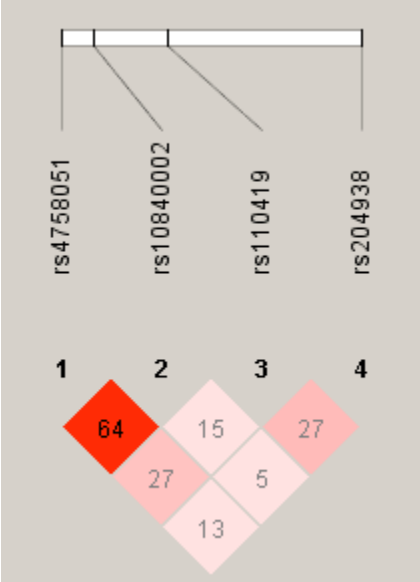


Supplementary Figure 2. LD plot for the significant SNPs around the *LMO1* locus. The four SNPs are in moderate linkage disequilibrium with each other.

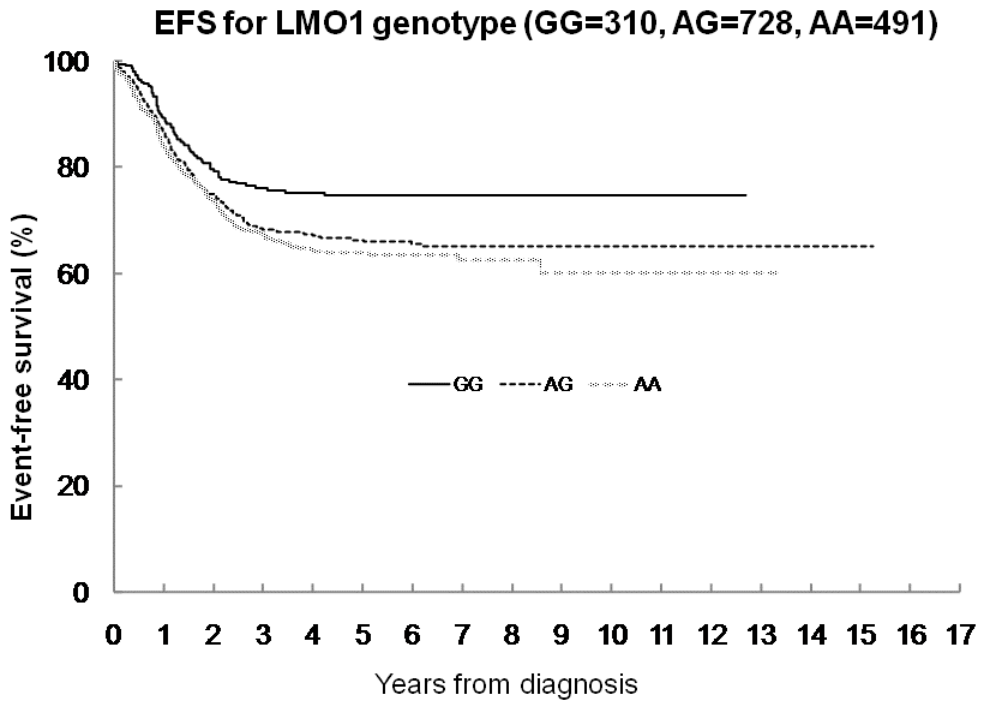
(a) D' measure:



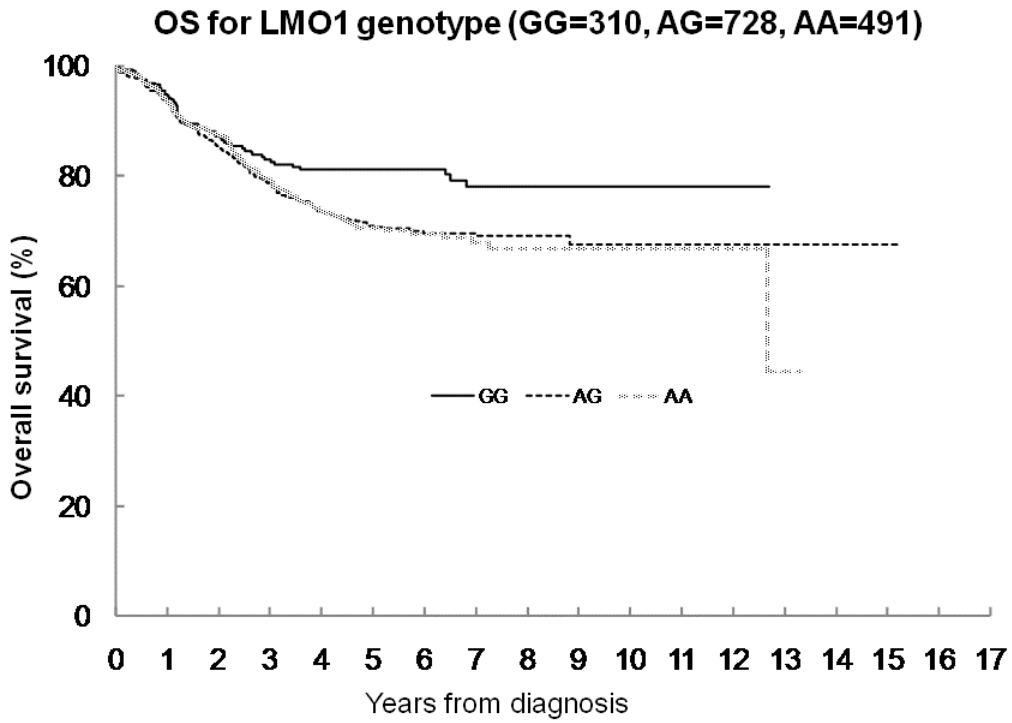
(b) r^2 measure



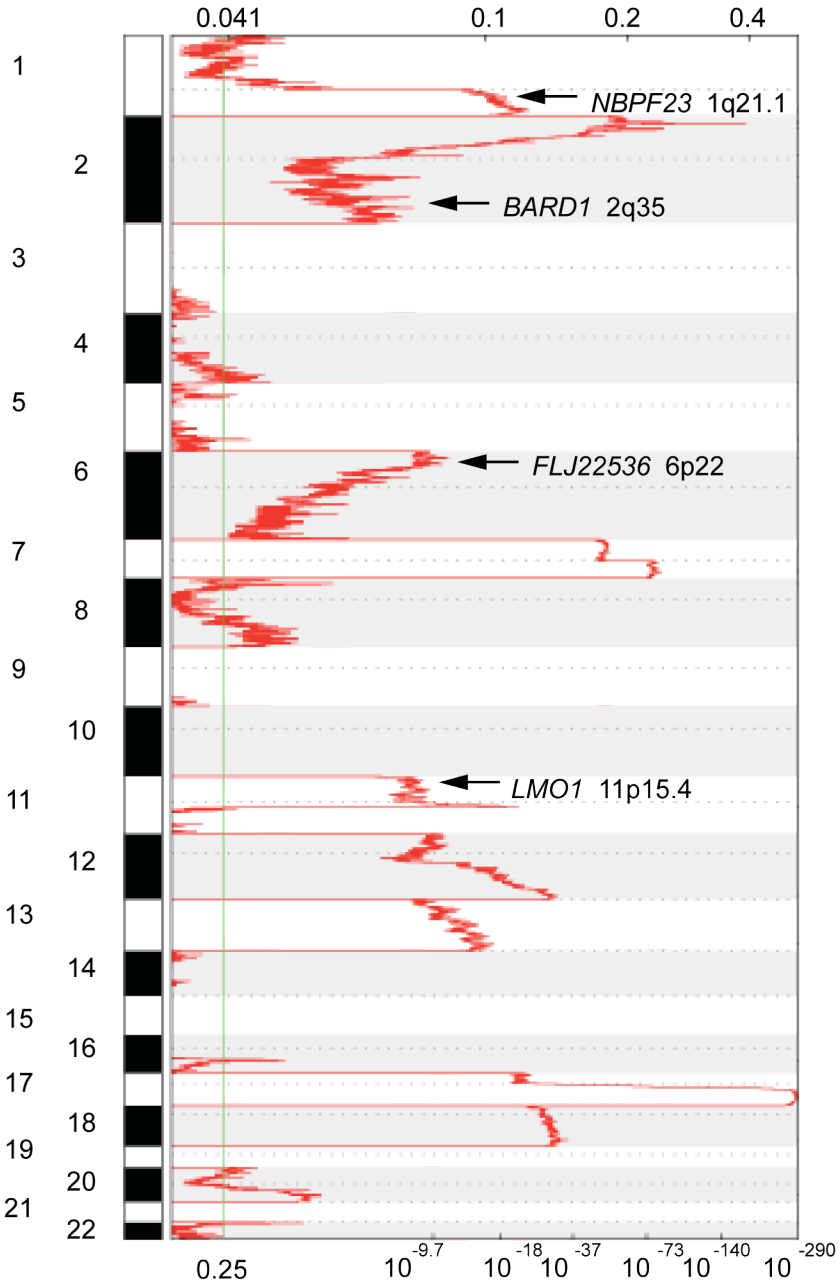
Supplementary Figure 3. Event-free survival for patients stratified by genotype of rs110419. The X-axis represents year from diagnosis. The AA, AG and GG genotypes represent homozygous risk genotype, heterozygous risk genotype and non-risk genotype, respectively. Patients carrying risk alleles have decreased event-free survival probability (P=0.0085)



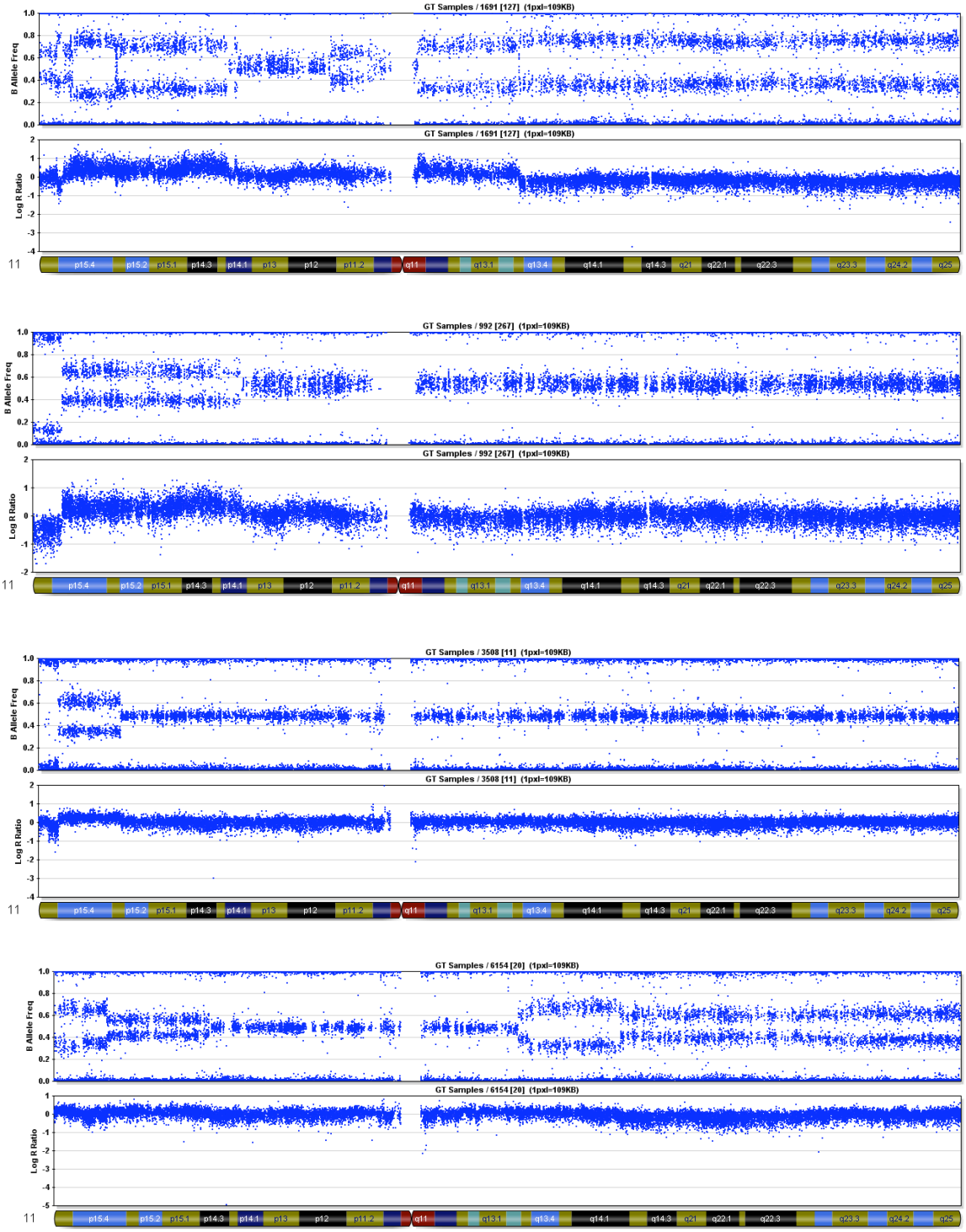
Supplementary Figure 4. Overall survival for patients stratified by genotype of rs110419. The X-axis represents year from diagnosis. The AA, AG and GG genotypes represent homozygous risk genotype, heterozygous risk genotype and non-risk genotype, respectively. Patients carrying the risk alleles have decreased overall survival probability (P=0.02).



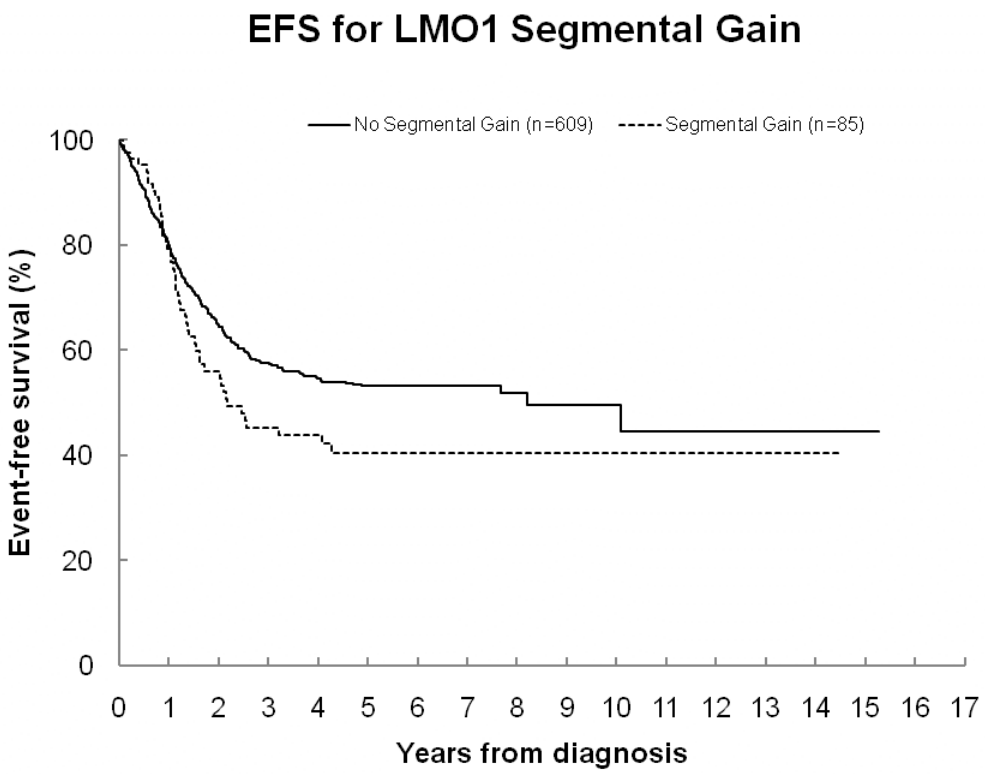
Supplementary Figure 5A . GISTIC [17] analysis of recurrent somatic copy number gains in 474 high-risk neuroblastoma tumors. Neuroblastoma susceptibility loci identified through our ongoing GWAS are annotated, including likely candidate genes. *LMO1* maps within a statistically significant broad region of gain ($q = 6.21 \times 10^{-8}$).



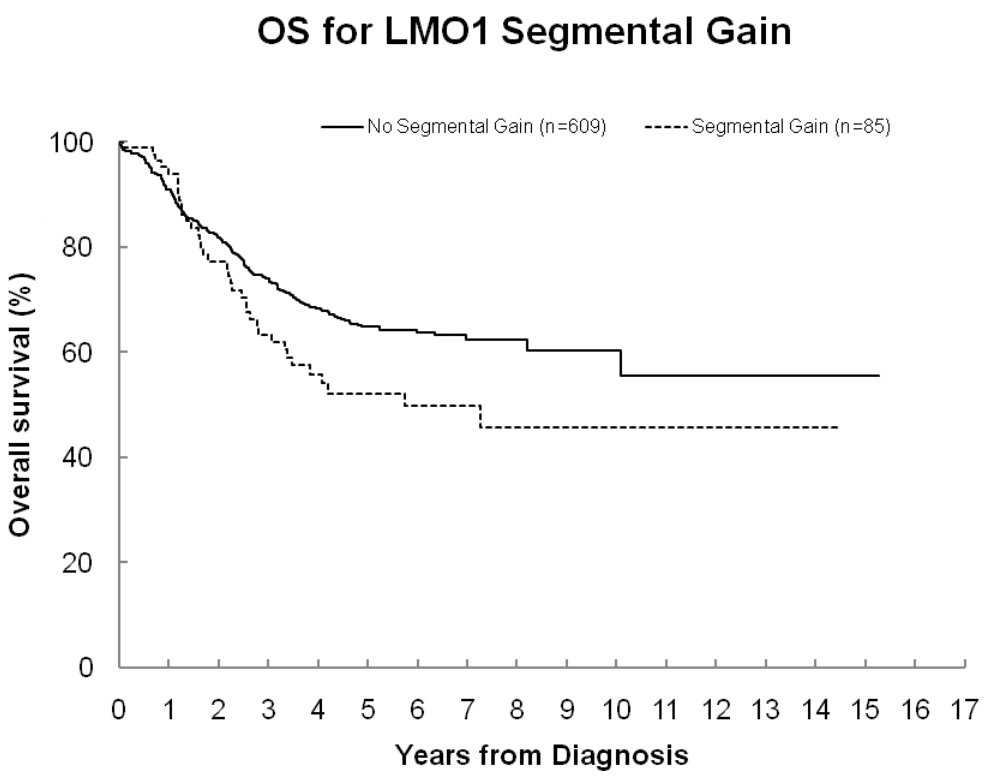
Supplementary Figure 5B. Four examples where additional focal copy number gains are present encompassing the *LMO1* locus on 11p15.4. The size of the focal gain is 15.8Mb, 15.6Mb, 9Mb and 15.5Mb, respectively.



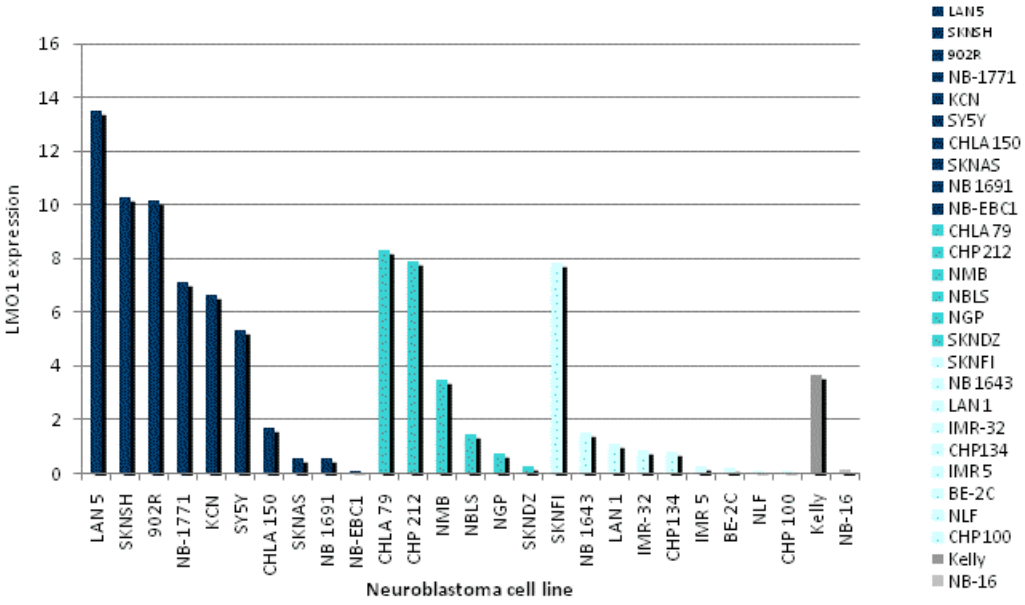
Supplementary Figure 6. Patients whose primary tumors carry *LMO1* gain have decreased event-free survival probability, but this observation did not reach statistical significance (P=0.077).

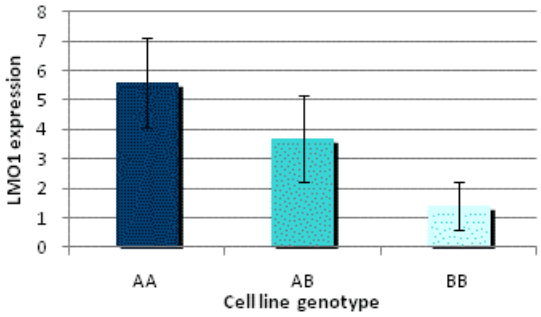


Supplementary Figure 7. Patients whose primary tumors carry *LMO1* gain have decreased overall survival probability (P=0.041).



Supplementary Figure 8. An expanded analysis on 27 cell lines to determine the correlation between rs110419 genotypes and *LMO1* expression levels by quantitative RT-PCR (these include the 9 cell lines shown in Figure 2A of the main manuscript). Error bars represent standard error of measurements. Two cell lines do not have genotype calls on rs110419 (grey bars).





11 References

1 Breslow, N. & McCann, B. Statistical estimation of prognosis for children with neuroblastoma. *Cancer Res* **31**, 2098-2103 (1971).

2 Brodeur, G. M. *et al.* Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. *Journal of Clinical Oncology* **11**, 1466-1477 (1993).

3 Shimada, H. *et al.* The International Neuroblastoma Pathology Classification (the Shimada system). *Cancer* **86**, 364-372 (1999).

4 Peuchmaur, M. *et al.* Revision of the International Neuroblastoma Pathology Classification: confirmation of favorable and unfavorable prognostic subsets in ganglioneuroblastoma, nodular. *Cancer* **98**, 2274-2281 (2003).

5 Look, A. T. *et al.* Clinical relevance of tumor cell ploidy and N-myc gene amplification in childhood neuroblastoma: a Pediatric Oncology Group study. *J Clin Oncol* **9**, 581-591 (1991).

6 Brodeur, G. M., Seeger, R. C., Schwab, M., Varmus, H. E. & Bishop, J. M. Amplification of N-myc in untreated human neuroblastomas correlates with advanced disease stage. *Science* **224**, 1121-1124 (1984).

7 Seeger, R. C. *et al.* Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. *N. Engl. J. Med.* **313**, 1111-1116 (1985).

8 Brodeur, G. M. & Maris, J. M. in *Principles and Practice of Pediatric Oncology* eds P.A. Pizzo & D.G. Poplack) 933-970 (J. B. Lippincott Company, 2006).

9 Maris, J. M. *et al.* Chromosome 6p22 locus associated with clinically aggressive neuroblastoma. *N. Engl. J. Med.* **358**, 2585-2593 (2008).

10 Capasso, M. *et al.* Common variations in BARD1 influence susceptibility to high-risk neuroblastoma. *Nat. Genet.* **41**, 718-723 (2009).

11 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559-575 (2007).

12 Gunderson, K. L., Steemers, F. J., Lee, G., Mendoza, L. G. & Chee, M. S. A genome-wide scalable SNP genotyping assay using microarray technology. *Nat Genet* **37**, 549-554 (2005).

13 Steemers, F. J. *et al.* Whole-genome genotyping with the single-base extension assay. *Nat Methods* **3**, 31-33 (2006).

14 Hakonarson, H. *et al.* A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. *Nature* **448**, 591-594 (2007).

15 Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904-909 (2006).

16 Attiyeh, E. F. *et al.* Genomic copy number determination in cancer cells from single nucleotide polymorphism microarrays based on quantitative genotyping corrected for aneuploidy. *Genome Res* **19**, 276-283 (2009).

17 Beroukhim, R. *et al.* Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *Proc. Natl. Acac. Sci. U S A* **104**, 20007-20012 (2007).

18 Diskin, S. J. *et al.* A copy number variation-based genome wide association study identifies NME7 as a predisposition locus and putative oncogene in neuroblastoma *Submitted* (2010).

19 Peto, R. *et al.* Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *Br J Cancer* **35**, 1-39 (1977).
20 Wang, Q. *et al.* Integrative genomics identifies distinct molecular classes of neuroblastoma and shows that multiple genes are targeted by regional alterations in DNA copy number. *Cancer Res* **66**, 6050-6062 (2006).
21 Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4**, 249-264 (2003).
22 Mosse, Y. P. *et al.* Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature* **455**, 930-935 (2008).
23 Cole, K. A. *et al.* A functional screen identifies miR-34a as a candidate neuroblastoma tumor suppressor gene. *Mol Cancer Res* **6**, 735-742 (2008).
24 Balamuth, N. J. *et al.* Serial transcriptome analysis and cross-species integration identifies centromere-associated protein E as a novel neuroblastoma target. *Cancer Res* **70**, 2749-2758 (2010).
25 Saeki, N. *et al.* GASDERMIN, suppressed frequently in gastric cancer, is a target of LMO1 in TGF-beta-dependent apoptotic signalling. *Oncogene* **26**, 6488-6498 (2007).